Commuting Extensions and Cubature Formulae

Ilan Degani

Departments of Mathematics and Chemical Physics

Weizmann Institute of Science, Rehovot 76100, Israel

email: `ilan.degani@weizmann.ac.il`

Jeremy Schiff

Department of Mathematics

Bar-Ilan University, Ramat Gan 52900, Israel

email: `schiff@math.biu.ac.il`

David Tannor

Department of Chemical Physics

Weizmann Institute of Science, Rehovot 76100, Israel

email: `david.tannor@weizmann.ac.il`

July 2004

Abstract

Based on a novel point of view on 1-dimensional Gaussian quadrature, we present a new approach to the computation of d-dimensional cubature formulae. It is well known that the nodes of 1-dimensional Gaussian quadrature can be computed as eigenvalues of the so-called Jacobi matrix. The d-dimensional analog is that cubature nodes can be obtained from the eigenvalues of certain mutually commuting matrices. These are obtained by extending (adding rows and columns to) certain noncommuting matrices $A_1, \ldots, A_d$, related to the coordinate operators $x_1, \ldots, x_d$, in $\mathbf{R}^d$. We prove a correspondence between cubature formulae and "commuting extensions" of $A_1, \ldots, A_d$, satisfying a compatibility condition which, in appropriate coordinates, constrains certain blocks in the extended matrices to be zero. Thus, the problem of finding cubature formulae can be transformed to the problem of computing (and then simultaneously diagonalizing) commuting extensions. We give a general discussion of existence and of the expected size of commuting extensions and describe our attempts at computing them, as well as examples of cubature formulae obtained using the new approach.

1 Introduction

One of the most elegant topics in numerical analysis is the theory of Gaussian quadrature [1]. Unfortunately this theory is limited to one dimension, and although something is known about generalizations to multiple dimensions (see [2] for a survey article and many references), at the moment there are many more questions than answers. The aim of this paper is to present a new approach to cubature rules ("cubature" seems to be the name given to the generalization of quadrature to arbitrary dimension). In classical, one-dimensional Gaussian quadrature, the most widely used method for computing nodes and weights, developed about 35 years ago [3], involves solving the eigenproblem for a certain tridiagonal matrix (see [4] for a recent "basis independent" discussion of this). As far as we are aware, no extension of this to higher dimensions has previously been obtained. Our proposed method for computing d-dimensional cubature formulae involves the construction of d matrices (with tridiagonal block structure in suitable bases), extending these, in a manner we will explain below, to a set of commuting matrices, and then solving the simultaneous eigenproblem for these commuting matrices.

The main novel step in this process, which follows very naturally from a new approach we present to one-dimensional Gaussian quadrature, is the need to construct *commuting extensions* of a set of matrices. We say the $N \times N$ matrices $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_d$ are $N \times N$ commuting extensions of the $n \times n$ matrices $A_1, A_2, \ldots, A_d$ (here $N \geq n$) if the top left $n \times n$ block in $\tilde{A}_i$ is A_i, for each $i = 1, 2, \ldots, d$, and the matrices $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_d$ pairwise commute. The idea of commuting extensions is very natural, but we do not find any such notion in the linear algebra or numerical linear algebra literature (see, however, some very similar ideas in a recent, independent, poster of Cargo and Littlejohn [5]). Since we hope the idea of commuting extensions will find other applications, the first few sections in this paper explore this subject without reference to cubature rules. Section 2 covers basic theory, section 3 discusses a couple of simple algorithms for computing commuting extensions, with which we currently have very limited success, and section 4 discusses commuting extensions when the matrices A_i take a special form, relevant for the study of cubature rules.

In section 5 we turn to the theory of cubature rules. Subsection 5.1 contains a novel approach to one dimensional Gaussian quadrature, based upon the properties of a certain operator, its eigenvalues and eigenfunctions. This serves as the model for all the subsequent discussion. In subsection 5.2 we consider the natural extension of this approach to multiple dimensions, and prove the central results of the paper, giving an equivalence between odd degree, positive weight cubature rules and commuting extensions (satisying a compatibility condition that will be explained in the sequel) of a certain set of matrices. Subsection 5.3

includes some simple consequences of this relationship. One of the key results in the theory of cubature rules, a lower bound on the number of nodes needed for an odd degree cubature rule, originating in the work of Möller [6], follows from a general result in the theory of commuting extensions (theorem 2 in section 3). Similarly a simple result on the spectra of commuting extensions (theorem 6 in section 3) gives interesting constraints on the nodes in positive weight cubature rules, which we believe have hitherto been overlooked even in one dimension.

In section 6 we turn to actual application of the commuting extension approach for computing nodes and weights. As in section 3, our achievements here are rather limited, nevertheless they validate our approach, and even give a few new cubature formulae. Section 7 contains a list of open questions.

We close this introduction by mentioning that one of the oldest continuing applications of Gaussian quadrature is in quantum mechanics, where it is used, in the so-called "DVR method" for the computation of matrix elements of non-exactly solvable Hamiltonians (see [7] for early references and [8] for recent reviews). This paper was born out of an attempt to extend the DVR method to higher dimensions; other, recent progress on this subject has been made by Dawes and Carrington [9]. Another interesting perspective on DVR can be found in the paper [10].

2 Commuting Extensions

In this section we present the basic theory of commuting extensions.

Definition. We say the $N \times N$ matrices $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_d$ are $N \times N$ *commuting extensions* of the $n \times n$ matrices $A_1, A_2, \ldots, A_d$ (here $N \geq n$) if the top left $n \times n$ block in $\tilde{A}_i$ is A_i, for each $i = 1, 2, \ldots, d$, and the matrices $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_d$ pairwise commute.

Theorem 1. Any set of matrices admits commuting extensions.

Proof: We construct explicit commuting extensions of the $n \times n$ matrices $A_1, A_2, \ldots, A_d$. Take

$$\tilde{A}_1 = \begin{pmatrix} A_1 & A_2 & A_3 & \ldots & A_d \\ A_d & A_1 & A_2 & \ldots & A_{d-1} \\ A_{d-1} & A_d & A_1 & \ldots & A_{d-2} \\ \vdots & \vdots & \vdots & & \vdots \\ A_2 & A_3 & A_4 & \ldots & A_1 \end{pmatrix}, \quad \tilde{A}_2 = \begin{pmatrix} A_2 & A_3 & A_4 & \ldots & A_1 \\ A_1 & A_2 & A_3 & \ldots & A_d \\ A_d & A_1 & A_2 & \ldots & A_{d-1} \\ \vdots & \vdots & \vdots & & \vdots \\ A_3 & A_4 & A_5 & \ldots & A_2 \end{pmatrix}, \quad \text{etc.} \tag{1}$$

More fully, take $\tilde{A}_i$ to be a $dn \times dn$ matrix which is a $d \times d$ matrix of $n \times n$ blocks, with the j,kth block, which we denote $(\tilde{A}_i)_{jk}$, equal to $A_{i+k-j \pmod d}$. Then

$$(\tilde{A}_i \tilde{A}_{i'})_{jj'} = \sum_{k=1}^{d} (\tilde{A}_i)_{jk} (\tilde{A}_{i'})_{kj'} = \sum_{k=1}^{d} A_{i+k-j \pmod d} A_{i'+j'-k \pmod d} \ . \tag{2}$$

Replacing the summation index k by $k' + i' - i \pmod d$ the second sum becomes

$$\sum_{k'=1}^{d} A_{i'+k'-j \pmod d} A_{i+j'-k' \pmod d} \tag{3}$$

which is equal to $(\tilde{A}_{i'} \tilde{A}_i)_{jj'}$. Thus

$$(\tilde{A}_i \tilde{A}_{i'})_{jj'} = (\tilde{A}_{i'} \tilde{A}_i)_{jj'} \tag{4}$$

for all j, j', i.e. $\tilde{A}_i$ and $\tilde{A}_{i'}$ commute. •

Theorem 1 establishes the existence of commuting extensions, but it is natural to ask what is the smallest possible dimension for commuting extensions of a given set of matrices. To this end we have the following result:

Theorem 2. If $N \times N$ commuting extensions of the $n \times n$ matrices $A_1, A_2, \ldots, A_d$ exist, then

$$N \geq n + \frac{1}{2} \mathrm{max}_{i,j} \mathrm{rank}([A_i, A_j]) \ . \tag{5}$$

Proof: Suppose $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_d$ are $N \times N$ commuting extensions of the $n \times n$ matrices $A_1, A_2, \ldots, A_d$. Write

$$\tilde{A}_i = \begin{pmatrix} A_i & a_i \\ b_i & \alpha_i \end{pmatrix} \ , \tag{6}$$

where the matrices a_i, b_i, α_i have sizes $n \times (N-n)$, $(N-n) \times n$, $(N-n) \times (N-n)$ respectively. The top left $n \times n$ block of the equation $[\tilde{A}_i, \tilde{A}_j] = 0$ gives the requirement

$$[A_i, A_j] + a_i b_j - a_j b_i = 0 \ . \tag{7}$$

Since the matrices a_i and b_i do not have rank exceeding $(N - n)$, neither do products of the form $a_i b_j$, and the matrices $a_i b_j - a_j b_i$ can have rank at most $2(N - n)$. Thus (7) can hold only if for each i, j we have

$$\mathrm{rank}([A_i, A_j]) \leq 2(N - n) \ , \tag{8}$$

and the theorem follows directly. •

Unfortunately there is a large gap between the lower bound on N from theorem 2 and the N in the existence proof of theorem 1. In practice, it seems that the lower bound of

theorem 2 is rarely attained, and the N of theorem 1 is much too big. As we shall see in Section 5, theorem 2 gives rise to a well known lower bound on the number of points needed for a cubature formula, and in that context also the bound can rarely be attained.

In addition to not knowing, in general, any way to rigorously predict the lowest dimension for commuting extensions of a given set of matrices, we also currently have no way of determining how many distinct families of commuting extensions of a given dimension exist. By a family we mean a set of commuting extensions related by conjugation as described in the following obvious result:

Theorem 3. If the matrices $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_d$ are $N \times N$ commuting extensions of the $n \times n$ matrices $A_1, A_2, \ldots, A_d$, then so are the matrices $\tilde{U}\tilde{A}_1\tilde{U}^{-1}, \tilde{U}\tilde{A}_2\tilde{U}^{-1}, \ldots, \tilde{U}\tilde{A}_d\tilde{U}^{-1}$, where $\tilde{U}$ is any matrix of the form

$$\tilde{U} = \begin{pmatrix} I_{n\times n} & 0_{n\times(N-n)} \\ 0_{(N-n)\times n} & U \end{pmatrix} \tag{9}$$

with U an invertible $(N-n) \times (N-n)$ matrix.

To proceed further, and at least get some idea of the size needed for commuting extensions, we have to resort to parameter counting. From here on we restrict to the case where the matrices A_i and $\tilde{A}_i$ are symmetric, i.e. the case of *symmetric* commuting extensions of a set of *symmetric* matrices. Note that except when $d = 2$ the existence construction of theorem 1 does not guarantee symmetric commuting extensions. Neither is it clear that the lowest dimension commuting extensions of a set of symmetric matrices need necessarily be symmetric. But because the case of symmetric commuting extensions of symmetric matrices is relevant for cubature rules, we restrict our attention to this.

If the matrices $\tilde{A}_i$ are symmetric then we can write

$$\tilde{A}_i = \begin{pmatrix} A_i & a_i \\ a_i^T & \alpha_i \end{pmatrix} , \tag{10}$$

where a_i is $n \times (N-n)$ and α_i is $(N-n) \times (N-n)$ and symmetric. Thus the number of free parameters we have in choosing the extensions of the A_i is

$$d\left(n(N-n) + \frac{1}{2}(N-n)(N-n+1)\right) = \frac{1}{2}d(N-n)(N+n+1) . \tag{11}$$

Let us assume that at least one of the $\tilde{A}_i$, say $\tilde{A}_1$, has distinct eigenvalues. Then all matrices that commute with $\tilde{A}_1$ also commute amongst themselves, and we just need to check that $[\tilde{A}_1, \tilde{A}_i] = 0$ for $i = 2, \ldots, d$. Since the commutator of symmetric matrices is automatically antisymmetric, we have

$$\frac{1}{2}N(N-1)(d-1) \tag{12}$$

equations to satisfy. We cannot, however, directly compare the number of parameters from (11) with the number of equations from (12), as from theorem 3 we learn that (except when $N = n+1$) commuting extensions exist in families. For symmetric commuting extensions the matrices U (and thus $\tilde{U}$) in theorem 3 are restricted to be orthogonal. So symmetric commuting extensions occur in families with $\frac{1}{2}(N-n)(N-n-1)$ parameters, and the number of parameters in choosing extensions should exceed the number of equations from (12) by at least this amount. Thus we need

$$\frac{1}{2}d(N-n)(N+n+1) \geq \frac{1}{2}N(N-1)(d-1) + \frac{1}{2}(N-n)(N-n-1) \ . \tag{13}$$

A little rearranging of this inequality gives the condition

$$N - n \geq \frac{n(n-1)(d-1)}{2(n+d)} = \frac{d-1}{2}n - \frac{d^2-1}{2} + \frac{d(d^2-1)}{2n} + o\left(\frac{1}{n}\right) \ . \tag{14}$$

If N satisfies this condition we expect to find $N \times N$ symmetric commuting extensions. For comparison, in the explicit commuting extensions of theorem 1 (which, however, were not symmetric) we had $N - n = (d-1)n$; we can clearly expect to do much better than this.

On occasions it seems that parameter counting can be misleading. As an example consider the case of $d = 2$. From theorem 2 we have $N - n \geq \frac{1}{2}\mathrm{rank}([A_1, A_2])$ (note that since the commutator of 2 symmetric matrices is antisymmetric, its rank is always even). The parameter counting argument tells us that we should expect

$$N - n \geq \frac{n(n-1)}{2(n+2)} \ . \tag{15}$$

Assuming $[A_1, A_2]$ of maximal rank we have

$$\frac{1}{2}\mathrm{rank}([A_1, A_2]) = \begin{cases} \frac{n}{2} & n \text{ even} \\ \frac{n-1}{2} & n \text{ odd} \end{cases} \ . \tag{16}$$

So by theorem 2

$$N - n \geq \begin{cases} \frac{n}{2} & n \text{ even} \\ \frac{n-1}{2} & n \text{ odd} \end{cases} \ . \tag{17}$$

We see that when $d = 2$ and $[A_1, A_2]$ is of maximal rank the inequality from parameter counting is actually weaker than the rigorous one from theorem 2. Thus in this case the parameter counting argument is certainly flawed. But this seems to be rather exceptional; in general it appears that when it is consistent with the lower bound of theorem 2, parameter counting gives a better idea of the size we should expect for commuting extensions. In particular we give the following example where the lower bound of theorem 2 cannot be attained, at least with symmetric extensions:

Theorem 4. For $n > 5$ there exist symmetric $n \times n$ matrices A_1, A_2 with $\mathrm{rank}\,([A_1, A_2]) = 2$ and no $(n+1) \times (n+1)$ symmetric commuting extensions.

The proof proceeds through the following lemma which will also be useful later:

Lemma 1. Let A_1, A_2 be a pair of symmetric $n \times n$ matrices with $\mathrm{rank}\,([A_1, A_2]) = 2$. Let $\{v, w\}$ be a basis of $\mathrm{Im}\,([A_1, A_2])$. If A_1, A_2 have $(n+1) \times (n+1)$ symmetric commuting extensions then the vectors $v, w, A_1v, A_1w, A_2v, A_2w$ are linearly dependent.

Proof (of lemma 1). $(n+1) \times (n+1)$ symmetric commuting extensions of A_1, A_2 must take the form

$$\tilde{A}_1 = \begin{pmatrix} A_1 & a \\ a^T & \alpha \end{pmatrix} , \qquad \tilde{A}_2 = \begin{pmatrix} A_2 & b \\ b^T & \beta \end{pmatrix} , \tag{18}$$

where a, b are n-dimensional column vectors and α, β are scalars. The requirement $[\tilde{A}_1, \tilde{A}_2] = 0$ translates into the equations

$$[A_1, A_2] + ab^T - ba^T \;=\; 0 \;, \tag{19}$$

$$A_1 b + \beta a - A_2 a - \alpha b \;=\; 0 \;. \tag{20}$$

Finding $(n+1) \times (n+1)$ symmetric commuting extensions of A_1, A_2 is equivalent to finding vectors a, b and scalars α, β satisfying (19)-(20).

Since $\mathrm{rank}\,([A_1, A_2]) = 2$, the vectors a, b cannot be linearly dependent, for if they were we would have $ab^T - ba^T = 0$, giving a contradiction with (19). Thus we can find a vector orthogonal to a but not to b. Applying (19) to this we deduce that a is in $\mathrm{Im}([A_1, A_2])$. Likewise, applying (19) to a vector orthogonal to b but not to a, we see that b is in $\mathrm{Im}([A_1, A_2])$. Choose $\{v, w\}$ to be a basis of $\mathrm{Im}\,([A_1, A_2])$. From the previous remarks it follows that we can write

$$a \;=\; \lambda v + \mu w \tag{21}$$

$$b \;=\; \nu v + \rho w \tag{22}$$

where λ, μ, ν, ρ are constants with $\lambda\rho - \mu\nu \neq 0$. Substituting in (20) we have

$$(\beta\lambda - \alpha\nu)v + (\beta\mu - \alpha\rho)w + \nu A_1 v + \rho A_1 w - \lambda A_2 v - \mu A_2 w = 0 \;. \tag{23}$$

Since $\lambda\rho - \mu\nu \neq 0$ we see at once that the 6 vectors $v, w, A_1v, A_1w, A_2v, A_2w$ must be linearly dependent. •

Proof (of Theorem 4). Let

$$(A_1)_{ab} = \begin{cases} \lambda_a & a = b \\ 0 & a \neq b \end{cases} , \qquad (A_2)_{ab} = \begin{cases} \mu_a & a = b \\ \frac{w_a v_b - w_b v_a}{\lambda_a - \lambda_b} & a \neq b \end{cases} , \tag{24}$$

where the λ_a are distinct, the μ_a are arbitrary, and the v_a, w_a are entries of two arbitrary linearly independent n-dimensional vectors. Then $[A_1, A_2] = wv^T - vw^T$, so rank $([A_1, A_2]) = 2$, and $\{v, w\}$ is a basis of Im $([A_1, A_2])$. By the lemma there can only exist $(n+1) \times (n+1)$ symmetric commuting extensions of A_1, A_2 if the 6 vectors $v, w, A_1 v, A_1 w, A_2 v, A_2 w$ are linearly dependent. There is no evident reason why they should be linearly dependent if $n > 5$, but to show concretely that they usually are not, we considered the case $n = 6$ where the λ_a take the values $1, 2, 3, 4, 5, 6$, and then constructed, using Maple, the determinant of the 6×6 matrix with columns $v, w, A_1 v, A_1 w, A_2 v, A_2 w$. The values of the v_a, w_a, μ_a were not fixed. The resulting determinant, which is too long to reproduce here, is simply a polynomial expression in the eighteen variables v_a, w_a, μ_a, and is not indentically zero. The case $n > 6$ follows trivially from the case $n = 6$. •

Note on index conventions: In discussion of commuting extensions we start with d matrices of size $n \times n$ which we extend to size $N \times N$. For clarity, in most of this paper we adhere to the following index conventions:
Indices i, j, k etc. run from 1 to d.
Indices a, b, c etc. run from 1 to n
Indices α, β, γ etc. run from 1 to N.

The results up to here all concern the existence and size of commuting extensions. For the purpose of finding commuting extensions, to be discussed in the next section, we will use the following:

Theorem 5. The $n \times n$ symmetric matrices $A_1, A_2, \ldots, A_d$ admit $N \times N$ symmetric commuting extensions if and only if there exist $N \times N$ diagonal matrices $\Lambda_1, \Lambda_2, \ldots, \Lambda_d$ and an $n \times N$ matrix Q with orthonormal rows such that

$$A_i = Q \Lambda_i Q^T \ . \tag{25}$$

Proof. *From the extensions to Λ_i, Q:* If we can find $N \times N$ symmetric commuting extensions $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_d$, then we can find diagonal matrices $\Lambda_1, \Lambda_2, \ldots, \Lambda_d$ and an $N \times N$ orthogonal matrix $\tilde{Q}$ such that

$$\tilde{A}_i = \tilde{Q} \Lambda_i \tilde{Q}^T \ . \tag{26}$$

The matrix Q in the theorem is comprised of just the first n rows of $\tilde{Q}$.

From Λ_i, Q to the extensions: A matrix Q as described in the theorem can always be extended, by the addition of $N - n$ orthonormal rows, to an $N \times N$ orthogonal matrix $\tilde{Q}$. (In fact this can be done in many ways, corresponding to the freedom described in Theorem 3.) Once such a $\tilde{Q}$ has been constructed the matrices $\tilde{A}_i = \tilde{Q} \Lambda_i \tilde{Q}^T$ are $N \times N$ symmetric commuting extensions of the A_i. •

Note: The matrix Q in theorem 5 satisfies $QQ^T = I_{n\times n}$.

It is of interest to understand how the spectra of commuting extensions (i.e. the entries of the matrices Λ_i in Theorem 5) are related to the spectra of the original matrices A_i. The following is a first result in this direction. It is a simple consequence of the Sturmian separation theorem [11], but we offer a direct proof too.

Theorem 6. Let $\tilde{A}$ be an $N \times N$ symmetric extension of the $n \times n$ matrix A. Then the smallest eigenvalue of $\tilde{A}$ is less than or equal to the smallest eigenvalue of A, and the largest eigenvalue of $\tilde{A}$ is greater than or equal to the largest eigenvalue of A.

Proof. Diagonalizing $\tilde{A}$ and A we have

$$\tilde{A} = \tilde{Q}\Lambda\tilde{Q}^T \ , \qquad A = UDU^T \ , \tag{27}$$

where Q and U are $N \times N$ and $n \times n$ orthogonal matrices, respectively, and Λ and D are $N \times N$ and $n \times n$ diagonal matrices containing the eigenvalues of $\tilde{A}$ and A, respectively. Restricting the first equation to its upper $n \times n$ block we have also

$$A = Q\Lambda Q^T \ , \tag{28}$$

where Q is an $n \times N$ matrix composed of the first n rows of $\tilde{Q}$, which are orthonormal. Comparing the two formulae for A and using the orthogonality of U we have

$$D = (U^TQ)\Lambda(U^TQ)^T \ . \tag{29}$$

The matrix U^TQ, like Q, has orthonormal rows; in particular we have

$$\sum_{\alpha=1}^{N}(U^TQ)^2_{a\alpha} = 1 \ , \qquad a = 1, \ldots, n \ . \tag{30}$$

The diagonal entries of (29) read

$$D_a = \sum_{\alpha=1}^{N}(U^TQ)^2_{a\alpha}\Lambda_\alpha \ , \qquad a = 1, \ldots, n \ . \tag{31}$$

Thus each eigenvalue D_a of A is a generalized average of the eigenvalues $\{\Lambda_\alpha\}$ of $\tilde{A}$ with respect to a set of positive weights $\{(U^TQ)^2_{a\alpha}\}$. It follows at once that it is not possible that any of the D_a be either smaller than, or greater than, all of the Λ_α. •

3 Computing Symmetric Commuting Extensions

The most obvious approach to computing commuting extensions is simply to treat the unknown entries in the extended matrices $\tilde{A}_i$ as variables, and to consider the conditions

$[\tilde{A}_i, \tilde{A}_j] = 0$ as equations in these variables. In the generic case (generically we should expect the $\tilde{A}_i$ have distinct eigenvalues) it will be sufficient to look at the equations just for one particular value of i. If $N - n > 1$, then by theorem 3 we expect continuous families of extensions, this freedom can be exploited to fix some of the variables. The system of equations we obtain will be quadratic in the unknown variables, and can be tackled by standard methods for solving systems of equations. We have done some initial experiments with this approach in the case $d = 2$, attempting to solve the system of quadratic equations 1) by integrating the gradient flow $v' = -\nabla ||[\tilde{A}_1(v), \tilde{A}_2(v)]||^2$ (here v denotes the variables added to form the commuting extensions), and 2) using Newton's method. The results are very varied; for some pairs of moderate-sized matrices there is reasonable convergence, but in other cases there are signs of extreme ill-coniditioning (very low gradients in the case of gradient flow, almost singular Jacobian in Newton's method). Some of the cubature related results we will present in section 6 were obtained by the integration of the gradient flow with Frobenius norm of the commutator; some technical details of these calculations can be found in [12].

In this section we focus on a different approach to computing commuting extensions, based on theorem 5 from section 2, and thus restricted to the case of symmetric extensions for symmetric matrices. We attempt to construct the matrices Λ_i and Q introduced in the theorem by minimization of

$$\begin{aligned} S(Q, \Lambda_1, \ldots, \Lambda_d) &= \frac{1}{2} \mathrm{Tr} \left(\sum_{i=1}^{d} (A_i - Q\Lambda_i Q^T)^2 \right) \\ &= \frac{1}{2} \sum_{i=1}^{d} \sum_{a,b=1}^{n} \left((A_i - Q\Lambda_i Q^T)_{ab} \right)^2 \ . \end{aligned} \tag{32}$$

Consider first the variation of this with respect to the entries of Λ_i (which is diagonal). It is straightforward to check that

$$\frac{\partial S}{\partial (\Lambda_i)_{\alpha\alpha}} = (Q^T (A_i - Q\Lambda_i Q^T) Q)_{\alpha\alpha} \qquad \begin{cases} i = 1, \ldots, d \\ \alpha = 1, \ldots, N \end{cases} , \tag{33}$$

implying that the entries of Λ_i must satisfy

$$\sum_{\beta=1}^{N} (Q^T Q)^2_{\alpha\beta} (\Lambda_i)_{\beta\beta} = (Q^T A_i Q)_{\alpha\alpha} \qquad \begin{cases} i = 1, \ldots, d \\ \alpha = 1, \ldots, N \end{cases} . \tag{34}$$

We see that if we can construct Q, we can, using the last formula, also construct Λ_i, at least assuming invertibility of the matrix with entries $(Q^T Q)^2_{\alpha\beta}$.

To build Q, or more precisely $\tilde{Q}$, the $N \times N$ orthogonal matrix that is an extension of the $n \times N$ matrix Q by addition of $(N - n)$ more rows, we use a sequence of Jacobi rotations

(not using the rotations that just mix the last $(N-n)$ rows). More fully: assuming we have an initial guess $\tilde{Q}_0$ for the matrix $\tilde{Q}$, we construct a new guess in the form $\tilde{Q} = R(\theta)\tilde{Q}_0$ where $R(\theta)$ is a Jacobi rotation of 2 specified rows of $\tilde{Q}_0$ through angle θ. θ is chosen to minimize S, where in S we use equation (34) to determine the Λ_i. The minimization is done numerically, there does not seem to be an explicit formula for θ available.

Note that this algorithm will actually solve a more general problem than that of finding commuting extensions. Suppose we use a value of N for which there are no commuting extensions. We can still compute a minimum of S, which means we will find Q and Λ_i such that $A_i \approx Q\Lambda_i Q^T$. Extending Q to an $N \times N$ orthogonal matrix $\tilde{Q}$ we will have commuting matrices $\tilde{A}_i = \tilde{Q}\Lambda_i\tilde{Q}^T$ which are *approximately* extensions of the A_i, i.e. we will be computing commuting approximate extensions. In the case $N = n$ we will be computing commuting approximations to the matrices A_i without increasing the dimension. The question of the existence and computation of commuting matrices which approximate a given set of matrices, with commutators of small norm, has a substantial mathematical history, having first been asked, apparently, in [13]. For some recent references see [14]. For the case of commuting approximations ($N = n$), the numerical approach we have just outlined was given in [15], though unlike in our more general case, it seems that in this case there is an explicit formula for the choice of rotation angle θ at each stage. This method is also advocated for numerical simultaneous diagonalization of commuting matrices. There are applications of the algorithm in statistics and in signal processing [16], and it is also employed in [9].

It remains to report on some numerical experiments. In all cases we implemented the algorithm starting with various random choices of $\tilde{Q}$ and applying sweeps of Jacobi rotations of all possible pairs of rows. We attempted to compute symmetric commuting extensions (and approximate extensions) for 4 specific pairs of symmetric 6×6 matrices; the findings are supported by many other numerical experiments as well. The specific matrices used can be found on the internet at

`http://www.math.biu.ac.il/~schiff/commext.html.`

The pairs of matrices chosen covered the following 4 cases:

1. $\mathrm{rank}([A_1, A_2]) = 2$, 7×7 symmetric commuting extensions known to exist.

2. $\mathrm{rank}([A_1, A_2]) = 2$, 7×7 symmetric commuting extensions known not to exist (see lemma 1, section 2).

3. $\mathrm{rank}([A_1, A_2]) = 4$, 8×8 symmetric commuting extensions known to exist, this being the lowest dimension allowed by theorem 2, section 2. (In fact presumably many

commuting extensions exist, as even after accounting for the invariance of theorem 3, section 2, there are more parameters than equations).

4. $\mathrm{rank}([A_1, A_2]) = 6$, 9×9 symmetric commuting extensions known to exist, this being the lowest dimension allowed by theorem 2, section 2. (Once again presumably many commuting extensions exist.)

In each case we not only ran the algorithm with N large enough that we expected to find commuting extensions, but also with all other $N \geq 6$ smaller than this (thus computing commuting approximations for $N = 6$ and commuting approximate extensions for $N > 6$ but too small for commuting extensions). Our observations can be summarized as follows:

1. For the actual computation of commuting extensions ($N = 7$ in case 1, $N = 8$ in cases 2 and 3, $N = 9$ in case 4) the positive result is that in each case the algorithm was observed to converge. After initial stabilization the value of S was observed to drop off according to

$$\ln S = a - bk \ , \tag{35}$$

where k is the sweep number, and a and $b > 0$ are (case-dependent) constants. This behavior was observed over many orders of magnitude. The problematic result is that *except in case 2* the value of b was found to be small, to the extent that obtaining even single precision accuracy required many thousands of sweeps. This behavior is not particularly surprising given that we are in practice doing a multidimensional optimization by sequential one-dimensional optimizations, with a fixed choice for the search directions. Remarkably in case 2 a large value of b was observed (and this persisted for other choices of matrices with rank 2 commutator but for which we were searching for extensions with $N = 8$).

2. For the computation of commuting approximations ($N = n$), in all cases the algorithm converged reasonably quickly, as reported in the literature [15], though there were noticeable differences in the rate of convergence, with case 1 being substantially slower than all the other cases. No problems with local minima were observed, different initial choices of $\tilde{Q}$ produced the same minimum value of S. Testing of this was limited, though.

3. In the search for commuting approximate extensions ($N = 7$ in cases 2 and 3, $N = 7, 8$ in case 4) a variety of behaviors were observed. For $N = 7$ in case 2 the search was, by far, the slowest performed, but it did ultimately converge. In case 4 the $N = 8$ search was noticeably slower than the $N = 7$ search, and for $N = 8$ we noticed convergence

to different values of S, indicating the presence of local (approximate) minima. In all cases the values of S achieved were at most a few orders of magnitude smaller than the values obtained when $N = n$. This gives hope that in the general case, when we do not know what the smallest dimension for commuting extensions is, we might be able to detect it by minimizing S for different N, but a better minimization algorithm than the current one will certainly be necessary.

It is clear from our results that a lot more work is necessary on the topic of computing commuting extensions.

4 A Special Case Of Commuting Extensions

We turn to consideration of a special case of commuting extensions that turns out to be relevant for cubature formula. It is characterized by 2 conditions: First, the symmetric $n \times n$ matrices A_i for which we wish to find commuting extensions have tridiagonal block form

$$A_i = \begin{pmatrix} \alpha_{i1} & a_{i1} & 0 & \dots & 0 & 0 \\ a_{i1}^T & \alpha_{i2} & a_{i2} & \dots & 0 & 0 \\ 0 & a_{i2}^T & \alpha_{i3} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{i(r-1)} & a_{i(r-1)} \\ 0 & 0 & 0 & \dots & a_{i(r-1)}^T & \alpha_{ir} \end{pmatrix} . \tag{36}$$

Here $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{ir}$ are symmetric square matrices of sizes $n_1 \times n_1, n_2 \times n_2, \dots, n_r \times n_r$ respectively, where $n_1 + n_2 + \dots + n_r = n$. The matrices $a_{i1}, a_{i2}, \dots, a_{i(r-1)}$ are of size $n_1 \times n_2, n_2 \times n_3, \dots, n_{r-1} \times n_r$ respectively. The second condition we impose is that the commutator matrices $[A_i, A_j]$ all vanish except for a single block in the bottom right hand corner, of size $n_r \times n_r$.

Let us seek symmetric commuting extensions with the matrices $\tilde{A}_i$, of size $N \times N$, also taking tridiagonal block form, that is

$$\tilde{A}_i = \begin{pmatrix} \alpha_{i1} & a_{i1} & 0 & \dots & 0 & 0 & 0 \\ a_{i1}^T & \alpha_{i2} & a_{i2} & \dots & 0 & 0 & 0 \\ 0 & a_{i2}^T & \alpha_{i3} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \alpha_{i(r-1)} & a_{i(r-1)} & 0 \\ 0 & 0 & 0 & \dots & a_{i(r-1)}^T & \alpha_{ir} & a_i \\ 0 & 0 & 0 & \dots & 0 & a_i^T & \alpha_i \end{pmatrix} , \tag{37}$$

where the new blocks α_i are of size $(N-n)\times(N-n)$ and are symmetric, and the a_i are of size $n_r\times(N-n)$.

The questions we wish to ask are (1) what are the equations that the new blocks α_i, a_i have to satisfy? and (2) how large need N be for us to have a hope that such extensions exist? As in section 2, we assume that $\tilde{A}_1$ has distinct eigenvalues, so we need only check that $\tilde{A}_1$ commutes with the $d-1$ matrices $\tilde{A}_2,\ldots,\tilde{A}_d$ and this guarantees that all the $\tilde{A}_i$ mutually commute. A brief calculation, using the fact that the commutators $[A_1, A_i]$ are zero except for a single block, gives the following conditions:

$$\left.\begin{array}{rcl} a_{1(r-1)}a_i - a_{i(r-1)}a_1 &=& 0 \\ a^T_{1(r-1)}a_{i(r-1)} - a^T_{i(r-1)}a_{1(r-1)} + \alpha_{1r}\alpha_{ir} - \alpha_{ir}\alpha_{1r} + a_1a_i^T - a_ia_1^T &=& 0 \\ \alpha_{1r}a_i - \alpha_{ir}a_1 + a_1\alpha_i - a_i\alpha_1 &=& 0 \\ a_1^Ta_i - a_i^Ta_1 + \alpha_1\alpha_i - \alpha_i\alpha_1 &=& 0 \end{array}\right\} \qquad i = 2,\ldots,d\ . \tag{38}$$

Of these four equations for each i, the first is of size $n_{r-1}\times(N-n)$, the second is of size $n_r\times n_r$ and antisymmetric, the third is of size $n_r\times(N-n)$ and the fourth is of size $(N-n)\times(N-n)$ and antisymmetric. Thus there is a total of

$$(d-1)\left(n_{r-1}(N-n) + \frac{1}{2}n_r(n_r-1) + n_r(N-n) + \frac{1}{2}(N-n)(N-n-1)\right) \tag{39}$$

equations to be satisfied. The number of variables available in the a_i and α_i is

$$d\left(n_r(N-n) + \frac{1}{2}(N-n)(N-n+1)\right)\ . \tag{40}$$

The system of equations (38), has an invariance

$$a_i \to a_i g\ , \qquad \alpha_i \to g^T\alpha_i g\ , \qquad i = 1,\ldots,r \tag{41}$$

where g is an $(N-n)\times(N-n)$ orthogonal matrix. Thus it is not sufficient that the number of variables simply exceed the number of equations to be solved, it must exceed the number of equations to be solved by at least $\frac{1}{2}(N-n)(N-n-1)$ to give a full family of solutions. Thus we can expect solutions provided:

$$\begin{aligned} d\left(n_r(N-n) + \tfrac{1}{2}(N-n)(N-n+1)\right) &\geq \tfrac{1}{2}(N-n)(N-n-1) \\ +(d-1)\left(n_{r-1}(N-n) + \tfrac{1}{2}n_r(n_r-1) + n_r(N-n) + \tfrac{1}{2}(N-n)(N-n-1)\right) & \end{aligned} \tag{42}$$

Simplifying this gives

$$N-n \geq \frac{n_r(n_r-1)}{2\left(\frac{n_r+d}{d-1} - n_{r-1}\right)}\ , \tag{43}$$

where we have made the assumption that the denominator on the right hand side is positive.

Note the right hand side in (43) does not depend on the total dimension, n, of the matrices A_i, but just on n_{r-1} and n_r. Thus the size of the extensions needed in this special case may well be substantially smaller than in general. The application to cubature of this special form of commuting extensions will become clear in section 5.3.

5 Cubature Formulae and Commuting Extensions

Much of the contents of this section, with some additions, are also discussed in [12]. In section 5.1 we give a novel presentation on the subject of Gaussian quadrature; in section 5.2 this is extended to the case of multidimensional cubature. Section 5.3 discusses some practical aspects and consequences of the results of 5.2.

5.1 A Novel Approach to Gaussian Quadrature

In Gaussian quadrature we wish to find $q+1$ nodes $x_0, \ldots, x_q$ and $q+1$ weights $w_0, \ldots, w_q$ such that the quadrature rule

$$\int_\Omega w(x)f(x)dx \approx \sum_{i=0}^{q} w_i f(x_i) \tag{44}$$

is exact whenever $f(x)$ is a polynomial of degree at most $2q+1$. Here Ω is some interval or union of intervals and $w(x) \geq 0$ is a suitable weight function. *Throughout this paper we only consider quadrature and cubature rules with positive weights, i.e.* $w_i > 0$.

Note: In this subsection there is no mention of commuting extensions so we allow ourselves to break our index conventions. In this subsection alone indices i, j, k run from 0 to q.

Denote by $\mathcal{P}_q$ the space of polynomials of degree at most q with the inner product

$$\langle a|b\rangle = \int_\Omega w(x)a(x)b(x)dx \ , \qquad \forall a, b \in \mathcal{P}_q \ . \tag{45}$$

Let Π_q be the projector from $\mathcal{P}_{q+1}$ onto $\mathcal{P}_q$ parallel to its orthogonal complement $\mathcal{P}_q^\perp$ *i.e.* the obvious orthogonal projection onto $\mathcal{P}_q$ with respect to the inner product above. We define the operator $\chi : \mathcal{P}_q \to \mathcal{P}_q$ by $\chi p = \Pi_q xp$ for all $p \in \mathcal{P}_q$. Since χ is self adjoint there is an orthonormal basis of $\mathcal{P}_q$ consisting of eigenfunctions $\{u_i\}$ of χ, $\chi u_i = \Lambda_i u_i$. Associated with χ there is a symmetric bilinear form X on $\mathcal{P}_q$ defined by $X(a,b) = \langle a|\chi b\rangle$, or equivalently

$$X(a,b) = \int_\Omega w(x)a(x)xb(x)\ dx \ , \qquad \forall a, b \in \mathcal{P}_q \ . \tag{46}$$

X is diagonalised in the basis $\{u_i\}$, $X(u_i, u_j) = \Lambda_i \delta_{ij}$.

We prove below that the eigenvalues $\{\Lambda_i\}$ of χ provide nodes for a Gaussian quadrature formula of degree $2q+1$. Our treatment is the reverse of the classical presentation of Gaussian quadrature [1], [3], see the explanation after theorem 8.

We need the following remarkable lemma:

Lemma 2.(The δ lemma) Let p be an arbitrary polynomial in $\mathcal{P}_{q+1}$. Then

$$\langle p|u_i\rangle = \langle 1|u_i\rangle p(\Lambda_i) \ , \tag{47}$$

i.e. the inner product of p with u_i is determined, up to normalization, by evaluation of p at Λ_i.

Note: In this lemma p is allowed to be in $\mathcal{P}_{q+1}$.

Proof. We prove recursively for j that

$$\langle x^j|u_i\rangle = \langle 1|u_i\rangle \Lambda_i^j \ , \qquad j = 0, \ldots, q+1 \ . \tag{48}$$

For $j = 0$ the statement is trivial. For $j > 0$,

$$\langle x^j|u_i\rangle = \langle \chi x^{j-1}|u_i\rangle = \langle x^{j-1}|\chi u_i\rangle = \Lambda_i \langle x^{j-1}|u_i\rangle \ . \tag{49}$$

This provides the recursive step proving (48), the full result follows by linearity. •

χ is an approximation of the operator x and this lemma shows that $\{u_i\}$, the eigenfunctions of χ, share with δ functions, which are "eigenfunctions" of x, the property that projection of a function on either is done by its evaluation at the appropriate eigenvalue. This similarity is the reason for the name we give to the δ lemma.

With this lemma it is almost immediate to prove the main result we want:

Theorem 7. Let f be a polynomial of degree at most $2q+1$. Then

$$\int_\Omega w(x) f(x) dx = \sum_{i=0}^{q} \langle 1|u_i\rangle^2 f(\Lambda_i) \ , \tag{50}$$

i.e. the quadrature rule

$$\int_\Omega w(x) f(x) dx \approx \sum_{i=0}^{q} \langle 1|u_i\rangle^2 f(\Lambda_i) \tag{51}$$

is exact of degree $2q+1$.

Proof. Again we prove the result for $f(x) = x^j$, $j = 0, \ldots, 2q+1$, the full result follows by linearity. For $j \geq 1$, choose integers n_1, n_2 between 0 and q such that $j = n_1 + n_2 + 1$. We then have

$$\int_\Omega w(x) x^j dx = X(x^{n_1}, x^{n_2}) = X\left(\sum_{k=0}^{q} \langle x^{n_1}|u_k\rangle u_k, \sum_{i=0}^{q} \langle x^{n_2}|u_i\rangle u_i\right) = \sum_{i=0}^{q} \langle 1|u_i\rangle^2 \Lambda_i^j \ . \tag{52}$$

In the last step we have used the δ lemma twice.

For $j = 0$ observe that

$$\int_\Omega w(x)dx = \langle 1|1\rangle = \left\langle \sum_{k=0}^{q} \langle 1|u_k\rangle u_k \middle| \sum_{i=0}^{q} \langle 1|u_i\rangle u_i \right\rangle = \sum_{i=0}^{q} \langle 1|u_i\rangle^2 \ , \tag{53}$$

by the orthonormality of the u_i. •

Theorem 7 relates the spectrum of χ to the nodes of Gaussian quadrature. It is in fact easy to prove other facts in the theory of Gaussian quadrature using our approach. For example, to see that none of the weights vanish just take $p = u_i$ in the δ lemma. To see that when Ω is a single interval $[a, b]$ the nodes must be in its interior of the interval just observe that

$$b - \Lambda_i = \int_a^b w(x)(b-x)u_i(x)^2 \ dx > 0, \qquad a - \Lambda_i = \int_a^b w(x)(a-x)u_i(x)^2 \ dx < 0 \ . \tag{54}$$

We also obtain the following widely known characterization of the nodes:

Theorem 8. The nodes Λ_i are roots of any nontrivial degree $q + 1$ polynomial orthogonal to $\mathcal{P}_q$.

Proof. Let p be a nontrivial degree $q + 1$ polynomial orthogonal to $\mathcal{P}_q$. Then using the δ lemma, $0 = \langle p|u_i\rangle = \langle 1|u_i\rangle p(\Lambda_i)$. Since $\langle 1|u_i\rangle$ is nonzero, this gives $p(\Lambda_i) = 0$. •

Our presentation on Gaussian quadrature is the reverse of that in [1], [3]. The starting point in [1], [3] is that the nodes in degree $2q + 1$ Gaussian quadrature are roots of the degree $q+1$ polynomial p from theorem 8; it is then shown that the eigenvalues of a matrix representation of χ are equal to these. Here we have gone in the other direction; without *a priori* assumption of the existence of a Gaussian quadrature formula, we have shown that the eigenvalues of χ are quadrature nodes and as a consequence of the δ lemma we also obtain the fact that they are roots of the degree $q + 1$ polynomial p from theorem 8.

As we shall see in the next subsection, our approach to Gaussian quadrature allows a generalization to higher dimensions. However, in the generalization of the δ lemma the polynomial p is restricted to $\mathcal{P}_q$, not $\mathcal{P}_{q+1}$, and this means that while theorem 7 can be generalized, theorem 8 cannot, at least not immediately. So the characterization of cubature nodes as roots of a polynomial (or a set of polynomials) seems to be lost. However the characterization of nodes as eigenvalues persists, as we now set out to show.

5.2 Generalization to Cubature

We denote by $\mathcal{P}_q$ the $n = \begin{pmatrix} d+q \\ d \end{pmatrix}$ dimensional vector space of polynomials in d variables $x_1, \ldots, x_d$ of total degree up to q (total degree is defined by degree$(x_1^{m_1} x_2^{m_2} \ldots x_d^{m_d}) = m_1 +$

$m_2 + \ldots + m_d$).

An N-point d-dimensional cubature formula

$$\int_\Omega w(x)f(x)d^dx \approx \sum_{\alpha=1}^{N} w_\alpha f(x_\alpha) \tag{55}$$

is said to be of degree D if it is exact whenever $f(x)$ is a polynomial of total degree at most D and non-exact for at least one polynomial of total degree $D+1$. Here Ω is a suitable region in $\mathbf{R}^d$ and $w(x) \geq 0$ a suitable weight function. The weights w_α are assumed positive.

We supply $\mathcal{P}_q$ with the inner product

$$\langle a|b\rangle = \int_\Omega w(x)a(x)b(x)\ d^dx\ , \qquad \forall\ a,\ b \in \mathcal{P}_q\ , \tag{56}$$

and define Π_q, the orthogonal projection operator from $\mathcal{P}_{q+1}$ onto $\mathcal{P}_q$ with respect to the above inner product, in the obvious way. We can then define d self adjoint operators $\chi_1, \ldots, \chi_d$ on $\mathcal{P}_q$ by

$$\chi_i p = \Pi_q x_i p\ , \qquad \forall p \in \mathcal{P}_q\ , \tag{57}$$

with related symmetric bilinear forms $X_i : \mathcal{P}_q \times \mathcal{P}_q \to \mathbf{R}$,

$$X_i(a,b) = \langle a|\chi_i b\rangle = \int_\Omega w(x)a(x)x_i b(x)\ d^dx \qquad \forall a, b \in \mathcal{P}_q\ . \tag{58}$$

Generally $[\chi_i,\ \chi_j] \neq 0$ so we can not find a basis of $\mathcal{P}_q$ in which all the χ_i are simultaneously diagonalised, and we do not have a direct analog of the one-dimensional case in which the eigenvalues of the single operator χ served as quadrature nodes. We shall show, however, that there is a correspondence between cubature rules and spectra of certain commuting extensions of matrix representations of the operators $\chi_1, \ldots,\ \chi_d$. As a first step towards this we prove the following:

Theorem 9. Let the $n \times n$ matrices $A_1, \ldots, A_d$ be the representations of the operators $\chi_1, \ldots, \chi_d$ in an arbitrary orthonormal basis $\{e_a\}$ of $\mathcal{P}_q$ (so $(A_i)_{ab} = \langle e_a|\chi_i e_b\rangle$). Suppose that for the region Ω and weight function $w(x)$ we have a degree $2q+1$, N point cubature rule with positive weights. Then there exist $N \times N$ symmetric commuting extensions of $A_1, \ldots, A_d$.

Proof. Suppose the cubature rule takes the form

$$\int_\Omega w(x)f(x)d^dx \approx \sum_{\alpha=1}^{N} w_\alpha f(x_\alpha)\ . \tag{59}$$

Then, since all integrands are of degree at most $2q+1$, we have

$$\begin{aligned} \delta_{ab} &= \langle e_a|e_b\rangle = \int_\Omega w(x)e_a(x)e_b(x)d^dx = \sum_{\alpha=1}^{N} w_\alpha e_a(x_\alpha)e_b(x_\alpha)\ , && (60) \\ (A_i)_{ab} &= \langle e_a|\chi_i e_b\rangle = \int_\Omega w(x)e_a(x)x_i e_b(x)d^dx = \sum_{\alpha=1}^{N} w_\alpha e_a(x_\alpha)(x_\alpha)_i e_b(x_\alpha)\ . && (61) \end{aligned}$$

Define the $n \times N$ matrix Q by

$$Q_{a\alpha} = \sqrt{w_\alpha} e_a(x_\alpha) \ , \tag{62}$$

and $N \times N$ diagonal matrices $\Lambda_1 \ \ldots, \Lambda_d$ with diagonal entries $(\Lambda_i)_{\alpha\alpha} = \Lambda_{i\alpha}$,

$$\Lambda_{i\alpha} = (x_\alpha)_i \ . \tag{63}$$

Equations (60)-(61) read

$$I_{n\times n} = QQ^T \ , \qquad A_i = Q\Lambda_i Q^T \ . \tag{64}$$

Using theorem 5 in section 2 we conclude that $A_1, \ldots, A_d$ have $N \times N$ symmetric commuting extensions. •

It is natural to ask whether the matrix commuting extensions of theorem 9 are representations of commuting extensions of the operators $\{\chi_i\}$ in some N dimensional space of functions V which includes $\mathcal{P}_q$ as a subspace. In other words, is it possible to extend the basis $\{e_a\}$ of $\mathcal{P}_q$ to an orthonormal basis of V by adding $N-n$ orthonormal functions $e_{n+1}, \ldots, e_N$ in such a way that the $N \times N$ matrices $(\tilde{A}_i)_{\alpha,\beta} = \langle e_\alpha | x_i e_\beta \rangle = \int_\Omega w(x) e_\alpha(x) x_i e_\beta(x) \ d^d x$ commute? Unfortunately, in all but the simplest cases, we could not find, nor prove the existence of, such functions $e_{n+1}, \ldots, e_N$. However, even though we can not view the matrix commuting extensions of theorem 9 as representations of operator extensions of the χ_i, they do satisfy a certain compatibility condition with the χ_i which we prove in theorem 10.

Let us introduce the N dimensional vector space $V(=\mathbf{R}^N)$, with the standard inner product, whose elements we denote in bold face. $\mathcal{P}_q$ is mapped to a subspace of V by the inclusion operator $\iota : \mathcal{P}_q \to V$ which is defined by $\iota e_1 = \mathbf{e_1}, \ldots, \iota e_n = \mathbf{e_n}$, where $\{\mathbf{e_a}\}$ are the first n members of the standard basis of V. Extend $\{\mathbf{e_a}\}$ to an orthonormal basis $\{\mathbf{e}_\alpha\}$ of V by adding any orthonormal basis $\{\mathbf{e_{n+1}}, \ldots, \mathbf{e_N}\}$ of the orthogonal complement of span$(\{\mathbf{e_a}\})$ in V. Even though our attempts to extend $\mathcal{P}_q$ with functions failed, now we are extending with N-tuples. Define the obvious projection operator $\pi : V \to \mathcal{P}_q$ by $\pi\mathbf{e_1} = e_1, \ldots, \pi\mathbf{e_n} = e_n$, $\pi\mathbf{e_{n+1}} = \ldots = \pi\mathbf{e_N} = 0 \in \mathcal{P}_q$; clearly $\pi\iota = I$, the identity operator on $\mathcal{P}_q$.

Recall that the essential step in the proof of theorem 5 is extension of Q to an $N \times N$ orthonormal matrix $\tilde{Q}$ by appending any $N-n$ orthonormal rows. In this way $\tilde{A}_1, \ldots, \tilde{A}_d$, $N \times N$ symmetric commuting extensions of the A_i are constructed in theorem 9, where $\tilde{A}_i = \tilde{Q}\Lambda_i\tilde{Q}^t$. Since the $\tilde{A}_1, \ldots, \tilde{A}_d$ mutually commute and are symmetric, there exist N orthonormal common eigenvectors $\mathbf{u}_\alpha \in V$, such that $\tilde{A}_i\mathbf{u}_\alpha = \Lambda_{i\alpha}\mathbf{u}_\alpha$. The matrices $\tilde{A}_i$ are given in the basis $\{\mathbf{e}_\alpha\}$ and $\tilde{Q}$ is the transformation between this basis and the eigenvector basis $\{\mathbf{u}_\alpha\}$. Note that the rows of $\tilde{Q}$ give the coordinates of the vectors $\{\mathbf{e}_\alpha\}$ in the basis

$\{\mathbf{u}_\alpha\}$, hence the extension of Q to $\tilde{Q}$ by adding arbitrary $N-n$ orthonormal rows is nothing but the extension of $\{\mathbf{e_a}\}$ to $\{\mathbf{e}_\alpha\}$ the orthonormal basis of V described above. The reader is reminded that at present we are assuming the existence of a cubature formula hence the eigenvalues $\Lambda_{i\alpha}$ are defined in (63). We can now state:

Theorem 10. The commuting extensions of theorem 9, $\{\tilde{A}_i\}$, satisfy the following compatibility condition with the operators $\{\chi_i\}$,

$$\tilde{A}_i\iota p = \iota x_i p = \iota\chi_i p, \qquad \forall p \in \mathcal{P}_{q-1} \ , \tag{65}$$

where $\mathcal{P}_{q-1}$ is regarded in the natural way as a subspace of $\mathcal{P}_q$.

Note: Applying π to (65) gives $\pi\tilde{A}_i\iota p = \chi_i p$, for $p \in \mathcal{P}_{q-1}$, which is automatic as $\tilde{A}_i$ is an extension of A_i. However, (65) contains more information than this, and is not true for an arbitrary extension of A_i.

Proof. We first prove that for any $p \in \mathcal{P}_q$ and any eigenvector $\mathbf{u}_\alpha$,

$$\langle \iota p | \mathbf{u}_\alpha \rangle = \sqrt{w_\alpha} p(x_\alpha) \ , \tag{66}$$

where w_α, x_α, are the weights and nodes of the cubature formula whose existence is assumed in theorem 9. We already noted that the rows of $\tilde{Q}$ from the proof of theorem 9 give the coordinates of the N vectors $\mathbf{e}_\alpha$ in the basis $\{\mathbf{u}_\alpha\}$, in particular the rows of Q give the coordinates of $\mathbf{e_a} = \iota e_a$, $a = 1, \ldots, n$. Recall that $Q_{a\alpha} = \sqrt{w_\alpha} e_a(x_\alpha)$, thus (66) is proven for basis elements $e_a \in \mathcal{P}_q$. The proof for general $p \in \mathcal{P}_q$ follows by linearity.

We now expand the left hand side of (65) in the basis $\{\mathbf{u}_\alpha\}$. For any $p \in \mathcal{P}_{q-1}$

$$\langle \tilde{A}_i \iota p | \mathbf{u}_\alpha \rangle = \Lambda_{i\alpha} \langle \iota p | \mathbf{u}_\alpha \rangle = \Lambda_{i\alpha} \sqrt{w_\alpha} p(x_\alpha) = \sqrt{w_\alpha} (x_\alpha)_i p(x_\alpha) \ , \tag{67}$$

using (66) and (63) respectively in the last two steps. To expand the right hand side of (65) note that $x_i p = \chi_i p \in \mathcal{P}_q$ for all $p \in \mathcal{P}_{q-1}$. Invoking (66) again we obtain

$$\langle \iota x_i p | \mathbf{u}_\alpha \rangle = \sqrt{w_\alpha} (x_\alpha)_i p(x_\alpha) \ , \tag{68}$$

which completes the proof. •

Note that taking $p = 1$ in (66) gives $w_\alpha = \langle \iota 1 | \mathbf{u}_\alpha \rangle^2$. In theorem 9 we saw that the eigenvalues of the commuting extensions are related to cubature nodes, here we obtain a relation between the weights and common eigenvectors.

We shall see in section 5.3 that with an appropriate choice of basis the compatibility condition of theorem 10 implies that the commuting extensions have certain off-diagonal zero blocks; in particular this special structure aids computation of commuting extensions.

The obvious question to ask at this stage is whether there is a converse of theorems 9 and 10. That is, suppose we have $\tilde{A}_1, \ldots, \tilde{A}_d$, $N \times N$ symmetric commuting extensions of $A_1, \ldots, A_d$, which satisfy the compatibility condition $\tilde{A}_i \iota p = \iota \chi_i p = \iota x_i p, \ \forall p \in \mathcal{P}_{q-1}$. Can we build a cubature rule, without *a priori* assumption of its existence, using the eigenvalues and eigenvectors of $\tilde{A}_1, \ldots, \tilde{A}_d$? In theorem 11 we give an affirmative answer to this. The treatment follows the presentation on Gaussian quadrature from section 5.1; in particular we start with a δ lemma. Note that given the commuting matrices $\tilde{A}_i$ we can find their diagonal representations Λ_i, but we do not assume in advance any connection of the Λ_i with cubature nodes.

Lemma 3 (The multidimensional δ lemma) Suppose the commuting extensions satisfy the compatibility condition $\tilde{A}_i \iota \grave{p} = \iota x_i \grave{p} = \iota \chi_i \grave{p}$ for all $i = 1, \ldots, d$, and for all $\grave{p} \in \mathcal{P}_{q-1}$. Then, for any $p \in \mathcal{P}_q$

$$\langle \iota p | \mathbf{u}_\alpha \rangle = \langle \iota 1 | \mathbf{u}_\alpha \rangle \ p(\lambda_\alpha) \ , \tag{69}$$

where the points $\lambda_\alpha \in \mathbf{R}^d$ have entries $(\Lambda_{1\alpha}, \ldots, \Lambda_{d\alpha})$, all eigenvalues of $\tilde{A}_1, \ldots, \tilde{A}_d$, satisfying $A_i \mathbf{u}_\alpha = \Lambda_{i\alpha} \mathbf{u}_\alpha$.

Proof. We prove (69) for monomials $p = x_1^{m_1} x_2^{m_2} \ldots x_d^{m_d}$. For $p = 1$ the statement is trivial. For any other monomial $p \in \mathcal{P}_q$ we can write $p = x_i \grave{p}$, where $\grave{p} \in \mathcal{P}_{q-1}$. Then

$$\langle \iota p | \mathbf{u}_\alpha \rangle = \langle \iota x_i \grave{p} | \mathbf{u}_\alpha \rangle = \langle \tilde{A}_i \iota \grave{p} | \mathbf{u}_\alpha \rangle = \langle \iota \grave{p} | \tilde{A}_i \mathbf{u}_\alpha \rangle = \Lambda_{i\alpha} \ \langle \iota \grave{p} | \mathbf{u}_\alpha \rangle \ . \tag{70}$$

Here the compatibility condition was used in the second step. Repeated application of (70) completes the proof for monomial p; the full result follows by linearity. •

Note: Recall that we do not know how to relate V to a space of polynomials (or other functions) in a way which gives commuting extensions of the operators $\chi_1, \ldots, \chi_d$. In particular, we can not interpret the eigenvectors $\mathbf{u}_\alpha$ as polynomials (or other functions). Thus our present state of understanding allows us to view the $\mathbf{u}_\alpha$ as "*δ vectors*" in V and not as δ *functions*, which was possible in the 1-dimensional case. Moreover we can not identify $\mathcal{P}_{q+1}$ with a subspace of V. Hence, in contrast to the one dimensional case, we restrict $p \in \mathcal{P}_q$ in the multidimensional δ lemma thereby losing the immediate connection between cubature nodes and roots of polynomials in $\mathcal{P}_q^\perp$.

We are now fully prepared for the converse statement to theorems 9 and 10:

Theorem 11. Let $A_1, \ldots, A_d$ be the representation of the operators $\chi_1, \ldots, \chi_d$ in an orthonormal basis $\{e_a\}$ of $\mathcal{P}_q$. Let $\tilde{A}_1, \ldots, \tilde{A}_d$ be $N \times N$ symmetric commuting extensions of $A_1, \ldots, A_d$ satisfying the compatibility condition $\tilde{A}_i \iota p = \iota x_i p = \iota \chi_i p$ for all $p \in \mathcal{P}_{q-1}$. Then

for every polynomial f in $\mathcal{P}_{2q+1}$

$$\int_\Omega w(x)f(x)d^dx = \sum_{\alpha=1}^{N} \langle \iota 1|\mathbf{u}_\alpha\rangle^2 f(\lambda_\alpha) \ . \tag{71}$$

Here the $\mathbf{u}_\alpha$ are joint eigenvectors of $\tilde{A}_1, \ldots, \tilde{A}_d$, satisfying $A_i\mathbf{u}_\alpha = \Lambda_{i\alpha}\mathbf{u}_\alpha$, and the points $\lambda_\alpha \in \mathbf{R}^d$ have entries $(\Lambda_{1\alpha}, \ldots, \Lambda_{d\alpha})$.

Proof. Recall the symmetric bilinear forms X_i on $\mathcal{P}_q$ defined in (58). Given the commuting extensions, we can introduce symmetric bilinear forms $\tilde{X}_i$ on V defined by $\tilde{X}_i(\mathbf{u}, \mathbf{v}) = \langle \mathbf{u}|\tilde{A}_i\mathbf{v}\rangle$. Since the $\tilde{A}_i$ are extensions of the A_i we have $X(p_1, p_2) = \tilde{X}_i(\iota p_1, \iota p_2)$ for all p_1, p_2 in $\mathcal{P}_q$. The $\tilde{X}_i$ are simultaneously diagonalized in the basis $\{\mathbf{u}_\alpha\}$, $\tilde{X}_i(\mathbf{u}_\alpha, \mathbf{u}_\beta) = \Lambda_{i\alpha}\delta_{\alpha\beta}$.

It is sufficient to prove the statement of the theorem for monomials. For $f = 1$ we have

$$\int_\Omega w(x)d^dx = \langle 1|1\rangle = \langle \iota 1|\iota 1\rangle = \left\langle \sum_{\alpha=1}^{N} \langle \iota 1|\mathbf{u}_\alpha\rangle \mathbf{u}_\alpha \middle| \sum_{\beta=1}^{N} \langle \iota 1|\mathbf{u}_\beta\rangle \mathbf{u}_\beta \right\rangle = \sum_{\alpha=1}^{N} \langle \iota 1|\mathbf{u}_\alpha\rangle^2 \ , \tag{72}$$

by the orthonormality of the $\mathbf{u}_\alpha$. Note that in the second expression in (72) the inner product is taken in $\mathcal{P}_q$, in subsequent expressions it is taken in V.

Any other monomial in $\mathcal{P}_{2q+1}$ can be written in the form $f = x_i f_1 f_2$ for some monomials $f_1, f_2 \in \mathcal{P}_q$ and some i. Note that use of the multidimensional δ lemma is possible since we assume the $\tilde{A}_i$ satisfy the compatibility condition, and so

$$\begin{aligned}
\int_\Omega w(x)f(x)d^dx &= \int_\Omega w(x)f_1(x)x_i f_2(x)\ d^dx = X_i(f_1, f_2) \\
&= \tilde{X}_i(\iota f_1, \iota f_2) = \tilde{X}_i\left(\sum_{\alpha=1}^{N} \langle \iota f_1|\mathbf{u}_\alpha\rangle \mathbf{u}_\alpha, \sum_{\beta=1}^{N} \langle \iota f_2|\mathbf{u}_\beta\rangle \mathbf{u}_\beta\right) \\
&= \sum_{\alpha=1}^{N} \Lambda_{i\alpha}\langle \iota f_1|\mathbf{u}_\alpha\rangle\langle \iota f_2|\mathbf{u}_\alpha\rangle = \sum_{\alpha=1}^{N} \langle \iota 1|\mathbf{u}_\alpha\rangle^2 \Lambda_{i\alpha} f_1(\lambda_\alpha) f_2(\lambda_\alpha) \\
&= \sum_{\alpha=1}^{N} \langle \iota 1|\mathbf{u}_\alpha\rangle^2 f(\lambda_\alpha) \ .
\end{aligned} \tag{73}$$

Thus (71) is proven for monomial f; the full result follows by linearity. •

Theorems 9,10,11 give the main result of this paper, that N point, odd order cubature formulae with positive weights are equivalent to symmetric commuting extensions, satisfying the compatibility condition, of matrix representations of the operators $\chi_1, \ldots, \chi_d$.

5.3 Discussion and Consequences

Our findings give a new computational approach to the derivation of cubature formulae. If appropriate commuting extensions are numerically found their simultaneous diagonalisation

will give the cubature rule in (71). To numerically obtain the matrices $\{A_i\}$ we introduce an orthonormal basis of $\mathcal{P}_q$ consisting of an orthonormal basis of $\mathcal{P}_0$ (a constant function e_1 with $\|e_1\| = \sqrt{\langle e_1|e_1\rangle} = 1$), extended to one of $\mathcal{P}_1$, extended to one of $\mathcal{P}_2$ etc., *i.e.* a basis $\{e_a\}$, $a = 1, \ldots, n = \begin{pmatrix} d+q \\ d \end{pmatrix}$, such that

$$\begin{aligned}
&e_1 \text{ is an orthonormal basis of } \mathcal{P}_0 \\
&e_1, \ldots, e_{d+1} \text{ is an orthonormal basis of } \mathcal{P}_1 \\
&e_1, \ldots, e_{\frac{1}{2}(d+1)(d+2)} \text{ is an orthonormal basis of } \mathcal{P}_2 \\
&\text{etc.}
\end{aligned} \tag{74}$$

Such a basis can be obtained from the monomials $\{x_1^{m_1} \ldots x_d^{m_d}\}$, $m_1 + \ldots + m_d \leq q$, by the Gram-Schmidt procedure. Note that all basis elements e_a of degree m or more are orthogonal to $\mathcal{P}_{m-1}$. This choice of basis and the fact that $\{A_i\}$ represent the operators $\{\chi_i\}$ imply that the A_i have tridiagonal block form as in (36), with $q+1$ blocks on the diagonal. In the notation of section 4, $n_1 = 1$ and for $m = 2, 3, \ldots, q+1$, $n_m = \dim\mathcal{P}_{m-1} - \dim\mathcal{P}_{m-2}$ ($n_1 = 1$, $n_2 = d$, $n_3 = d(d+1)/2$ etc). Moreover, the commutator of any pair of A_i's is zero apart from a single block of size $\begin{pmatrix} q+d-1 \\ d-1 \end{pmatrix} \times \begin{pmatrix} q+d-1 \\ d-1 \end{pmatrix}$ in the bottom right hand corner (note $\begin{pmatrix} q+d-1 \\ d-1 \end{pmatrix} = n_{q+1} = \dim\mathcal{P}_q - \dim\mathcal{P}_{q-1}$). The compatibility condition, together with the fact that $\mathbf{e_{n+1}}, \ldots, \mathbf{e_N}$ are orthogonal to $\iota\mathcal{P}_q$, imply that the bottom left $(N-n)\times\dim(\mathcal{P}_{q-1})$ block of $\tilde{A}_i$ is zero and by symmetry so is the corresponding block in the upper right. Thus the commuting extensions $\{\tilde{A}_i\}$ we seek are precisely those in tridiagonal block form as in section 4. Note also that since our first basis element e_1 is a constant polynomial the cubature weights are obtained from the first entries of the eigenvectors $\mathbf{u}_\alpha$, $w_\alpha = \langle \iota 1|\mathbf{u}_\alpha\rangle^2 = \frac{1}{e_1^2}\langle \iota e_1|\mathbf{u}_\alpha\rangle^2 = \left(\frac{(\mathbf{u}_\alpha)_1}{e_1}\right)^2$.

In the case $d = 2$, we note that the matrices A_1, A_2 have $n_{r-1} = q$ and $n_r = q+1$ in the notation of section 4, and thus from (43), which is based on counting degrees of freedom, the expected size of the commuting extensions is

$$N \geq n + \frac{q(q+1)}{6} \ . \tag{75}$$

Using $n = \dim\mathcal{P}_q = \frac{1}{2}(q+1)(q+2)$ we obtain

$$N \geq \frac{(2q+2)(2q+3)}{6} = \frac{1}{3}\dim\mathcal{P}_{2q+1} \ . \tag{76}$$

This is exactly the number of nodes we expect from counting degrees of freedom in a 2-dimensional cubature formula of degree $2q+1$.

For $d > 2$ the A_i have a more subtle structure that requires refinement of the discussion leading to equation (43). However, the equivalence of cubature formulae and commuting extensions (satisfying the compatibility condition) allows us to estimate N by easily counting degrees of freedom in a general d-dimensional cubature formula of degree $2q+1$. Thus,

$$N \geq \frac{1}{d+1} \dim \mathcal{P}_{2q+1} \ . \tag{77}$$

We emphasize again that such calculations are not rigorous and the inequalities obtained in this way can serve only as recomendations for choice of N, indeed certain cubature formulas with less points are known.

In section 6 we shall give some first examples of computation of cubature nodes using our approach. But before we do this we present two theoretical consequences of the equivalence between cubature formulae and commuting extensions.

Theorem 12. Let N be the number of nodes in a degree $2q+1$, d-dimensional positive weight cubature rule. Then

$$N \geq \binom{d+q}{d} + \frac{1}{2} \max_{i,j} \mathrm{rank}([A_i, A_j]) \ , \tag{78}$$

where $A_1, \ldots, A_d$ are the matrix representations of the operators $\chi_1, \ldots, \chi_d$ on $\mathcal{P}_q$.

Proof. By theorem 9 an N point cubature rule gives $N \times N$ commuting extensions of the matrices A_i. By theorem 2, section 2, the size of such extensions is at least $\dim \mathcal{P}_q + \frac{1}{2} \max_{i,j} \mathrm{rank}([A_i, A_j])$. •

Notes: (1) As mentioned in the introduction, theorem 12 has its origins in the work of Möller [6]. A statement of the result in a form that clearly corresponds to our statement can be found in [17], which cites [18] and [19]. Our proof, however, is a substantial simplification. (2) It is informative to compare the lower bound of theorem 12 with estimates based on parameter counting. As a consequence of our previous remarks on the block structure of $[A_i, A_j]$

$$\mathrm{rank}([A_i, A_j]) \leq \binom{d-1+q}{d-1} = \frac{d}{d+q} \binom{d+q}{d} \ , \tag{79}$$

so the second term in (78) is typically a small fraction of the first term. Consequently for large q the right hand side of (78) is much smaller than the number of nodes we expect from counting degrees of freedom in a degree $2q+1$, d-dimensional cubature formula, which is

$$\left\lceil \frac{1}{d+1} \binom{d+2q+1}{d} \right\rceil . \tag{80}$$

This comparison indicates why the lower bound on the number of points needed for a cubature formula is rarely attained.

Theorem 13. Let $A_1, \ldots, A_d$ be matrix representations of $\chi_1, \ldots, \chi_d$. In any degree $2q+1$, d-dimensional, positive weight cubature rule, and for each i, there is a node x_α with $(x_\alpha)_i$ less than or equal to the smallest eigenvalue of A_i, and a node x_β with $(x_\beta)_i$ greater than or equal to the largest eigenvalue of A_i.

Proof. By theorem 9 a cubature rule of degree $2q+1$ gives commuting extensions of the matrices A_i with the nodes composed of the eigenvalues of the extended matrices. By theorem 6 in section 2 the smallest/largest eigenvalue of the extended matrices is less/greater than or equal to the smallest/largest eigenvalue of the matrices before extension. $\bullet$

Note: As far as we are aware this theorem is not even known for $d = 1$. For $d = 1$ the theorem says that any N-point, positive weight, degree $2q+1$ quadrature rule must have a node less/greater than or equal to the smallest/largest Gaussian quadrature node. Thus Gaussian quadrature has the property that the span of the nodes is the smallest possible, amongst all positive weight quadrature rules, with any number of points, that are exact to the same degree.

6 Examples

In this section we briefly discuss 2 examples of computing cubature rules via commuting extensions, both are in 2 dimensions. In section 6.1 we consider the classic question, first studied by Radon [20], of finding 7 point cubature rules which are exact for polynomials up to degree 5 ($\dim \mathcal{P}_5 = 21$ for $d = 2$). This involves 7×7 extensions of a pair of 6×6 matrices of the tridiagonal block form of section 4, and we provide a reliable algorithm for this. In section 6.2 we present some new cubature rules for integration on the entire plane with weight function $w(x, y) = e^{-x^2-y^2}$. These were computed by commuting extension techniques, though, as explained in section 3, our current algorithms for this are poor, so we give limited details.

6.1 Radon type formulae

Given a region Ω on the plane and a suitable weight function $w(x_1, x_2)$, a Radon formula is a cubature rule of the form

$$\int_\Omega w(x_1, x_2) f(x_1, x_2)\ d^2x = \sum_{\alpha=1}^{7} w_\alpha f(x_{1\alpha}, x_{2\alpha}) \tag{81}$$

which is exact for all polynomials of degree no more than 5.

To construct such formulae we proceed as follows: First choose a basis $e_1, e_2, e_3, e_4, e_5, e_6$ of $\mathcal{P}_2$ which is orthonormal with respect to the inner product

$$\langle a|b\rangle = \int_\Omega w(x_1,x_2)a(x_1,x_2)b(x_1,x_2)\ d^2x\ , \tag{82}$$

and such that e_1 is of degree 0, e_2, e_3 are of degree 1, and e_4, e_5, e_6 are of degree 2. Typically we do this by applying Gram–Schmidt orthonormalization to the basis $1, x_1, x_2, x_1^2, x_1x_2, x_2^2$.

Having constructed the orthonormal basis we compute the matrices A_1 and A_2 via:

$$(A_1)_{ij} \ = \ \int_\Omega w(x_1,x_2)e_i(x_1,x_2)x_1e_j(x_1,x_2)\ d^2x\ , \tag{83}$$

$$(A_2)_{ij} \ = \ \int_\Omega w(x_1,x_2)e_i(x_1,x_2)x_2e_j(x_1,x_2)\ d^2x\ . \tag{84}$$

By orthogonality we will have $(A_1)_{1i} = (A_1)_{i1} = (A_2)_{1i} = (A_2)_{i1} = 0$ for $i = 4,5,6$ (this is the meaning of "tridiagonal block form" in this case), and the commutator $C = [A_1, A_2]$ will be all zero except for a single antisymmetric 3×3 block in the lower right hand corner. The commuting extensions $\tilde{A}_1, \tilde{A}_2$ should take the form given in (18), where a, b are 6-dimensional column vectors, with the first 3 entries vanishing (these are the $\dim\mathcal{P}_{q-1} \times (N-n)$ zero blocks described in section 5.3), and α, β scalars. Following the arguments in lemma 1 from section 2, a, b, α, β must satisfy (19)-(20). To solve these equations we proceed exactly as in the lemma. First we construct a specific pair of 6-dimensional vectors v, w satisfying

$$[A_1, A_2] + vw^T - wv^T = 0\ . \tag{85}$$

Because of the special structure of $[A_1, A_2]$, v, w can be taken to have zeros in their first 3 entries, and determining the other entries is equivalent to finding 2 vectors in $\mathbf{R}^3$ with a given cross product. Clearly v, w are linearly independent and give a basis for $\mathrm{Im}([A_1, A_2])$. Once such v, w have been found, (19) gives

$$a = \lambda v + \mu w\ , \qquad b = \nu v + \rho w\ , \tag{86}$$

where $\lambda\rho - \mu\nu = 1$, and (20) gives the requirement

$$(\beta\lambda - \alpha\nu)v + (\beta\mu - \alpha\rho)w + \nu A_1 v + \rho A_1 w - \lambda A_2 v - \mu A_2 w = 0\ , \tag{87}$$

or

$$\begin{pmatrix} v & w & A_1v & A_1w & A_2v & A_2w \end{pmatrix} \begin{pmatrix} \beta\lambda - \alpha\nu \\ \beta\mu - \alpha\rho \\ \nu \\ \rho \\ -\lambda \\ -\mu \end{pmatrix} = 0\ . \tag{88}$$

Since v, w have zeros in their first 3 entries, and $(A_1)_{1i} = (A_2)_{1i} = 0$ for $i = 4, 5, 6$, the entire first row of the matrix $(\, v \quad w \quad A_1 v \quad A_1 w \quad A_2 v \quad A_2 w \,)$ is zero. So it is singular and a nontrivial solution of $(\, v \quad w \quad A_1 v \quad A_1 w \quad A_2 v \quad A_2 w \,)\, k = 0$ is guaranteed. To complete construction of a commuting extension all we need is to find $\alpha, \beta, \lambda, \mu, \nu, \rho$ with $\lambda\rho - \mu\nu = 1$, and such that the column vector in (88) is in the kernel of the matrix in (88). Using the freedom to rescale vectors in the kernel, it is straightforward to conclude that given a nontrivial vector k in the kernel of the matrix, there will be an associated commuting extension if and only if $k_3 k_6 - k_4 k_5 > 0$, and we must take

$$\begin{gathered} \lambda = -ck_5 \ , \qquad \mu = -ck_6 \ , \qquad \nu = ck_3 \ , \qquad \rho = ck_4 \ , \\ \alpha = c^2(k_2 k_5 - k_1 k_6) \ , \qquad \beta = c^2(k_1 k_4 - k_2 k_3) \ , \\ \text{where} \quad c = \tfrac{1}{\sqrt{k_3 k_6 - k_4 k_5}} \end{gathered} \tag{89}$$

To summarize: to find a Radon formula one should: (1) Construct the orthonormal basis $\{e_a\}$. (2) Construct the matrices A_1, A_2. (3) Find 6-dimensional vectors v, w with top 3 entries zero, and such that $[A_1, A_2] + vw^T - wv^T = 0$. (4) Compute the kernel of the matrix $(\, v \quad w \quad A_1 v \quad A_1 w \quad A_2 v \quad A_2 w \,)$. Then for each vector k in the kernel with $k_3 k_6 - k_4 k_5 > 0$ there is an associated commuting extension given by (18), where a, b are given by (86) and $\alpha, \beta, \lambda, \mu, \nu, \rho$ by (89). The commuting extensions can be simultaneously diagonalized, using the algorithm in [15], to obtain the nodes and weights of the cubature rule.

This procedure should be compared with Radon's original work [20]. In practice we have not found a case in which there is a vector in the kernel with $k_3 k_6 - k_4 k_5 \leq 0$, but have no explanation why this is so. Generically the kernel is one dimensional and there is a unique Radon formula. For the case of Ω equal to the circle or the square each with uniform weight $w(x_1, x_2) = 1$, the kernel has dimension 2, giving a one parameter family of Radon formulas. In the case of the square with vertices $(-1,-1), (-1,1), (1,1), (1,-1)$, all the Radon formulae have a single node at the origin with weight $\frac{8}{7}$ and 3 pairs of diametrically opposed nodes on the circle $x_1^2 + x_2^2 = \frac{14}{15}$. (In the literature there is often only mention of the case of Radon formulae for the square when one pair of nodes lies on one of the coordinate axes.) In the case of the unit circle, there is a single node at the origin with weight $\frac{\pi}{4}$ and 3 pairs of diametrically opposed nodes on the circle $x_1^2 + x_2^2 = \frac{2}{3}$; in this case there is full rotational symmetry, and the different formulae are related by rotation.

As an explicit example of nonstandard Radon formulae, we give here some results for the square with vertices $(-1,-1), (-1,1), (1,1), (1,-1)$, with a square with vertices $(\frac{2}{5} - r, \frac{3}{5} - r), (\frac{2}{5} - r, \frac{3}{5} + r), (\frac{2}{5} + r, \frac{3}{5} + r), (\frac{2}{5} + r, \frac{3}{5} - r)$ removed. Here $0 \leq r \leq \frac{2}{5}$ and the weight is uniform. Figure 1 displays the location of the nodes for $r = \frac{i}{20}$, $i = 0, 1, \ldots, 8$,

along with the boundaries of the squares removed, and the circle $x_1^2 + x_2^2 = \frac{14}{15}$ on which the nodes lie for $r = 0$. For $r = 0$ we in fact plot the nodes for a very small nonzero value of r, so as to have a unique result. For values of r up to $\frac{1}{4}$ all the nodes are inside the domain of integration. For $r = \frac{3}{10}$ and $r = \frac{7}{20}$ one node lies inside the small square that is removed to obtain the cubature domain. For $r = \frac{2}{5}$ one node lies outside the larger square, at $(x_1, x_2) \approx (0.1844, 1.0360)$, but has low weight (about 3.25% of the sum of the weights).

Recall that our approach for constructing cubature formulae involves first constructing commuting extensions, and then diagonalizing them. For the rational values of r discussed in the previous paragraph, the commuting extensions can be constructed exactly, and we do not need to apply the numerical algorithms for commuting extensions from section 3 (the $S(Q, \Lambda_i)$ algorithm does not work particularly well for the matrices in this subsection). Even though the commuting extensions can be found exactly, finding the cubature nodes (the eigenvalues of the extended matrices) can only be done numerically.

6.2 Gaussian weight on the plane

As already explained, to construct degree $2q+1$ cubature formulae on a region Ω in the plane, we need to find symmetric commuting extensions (satisfying the necessary compatability conditions) of the $n \times n$ matrices defined in (83)-(84), where $n = \dim \mathcal{P}_q = (q+1)\left(\frac{1}{2}q + 1\right)$. By counting degrees of freedom, we expect N-dimensional symmetric extensions if $N \geq (q+1)\left(\frac{2}{3}q + 1\right)$, see the discussion leading to (76). We work in a basis of the type discussed in subsection 5.3, so the matrices are all of the tridiagonal block form of section 4. Using the notation of section 4, we extend each matrix by adding 2 blocks, a_i of size $(q+1) \times (N-n)$, and α_i (symmetric), of size $(N-n) \times (N-n)$, here $i = 1, 2$. Thanks to the freedom in choice of extensions described in theorem 3, one of the symmetric added blocks α_i can be chosen diagonal. Taking this into account we expect $N - n \geq \frac{1}{6}q(q+1)$. Even so, the number of entries added in the blocks $a_1, a_2, \alpha_1, \alpha_2$ grows as q^4.

For applications in quantum mechanics, cubature on the plane $\Omega = \mathbf{R}^2$ with weight function $w(x_1, x_2) = e^{-x_1^2 - x_2^2}$ is important. For this case we have computed degree 11, 13, 15 and 17 cubature formulas with 26, 35, 46 and 57 nodes respectively, by simultaneous diagonalization of commuting extensions. These were found using the gradient flow approach mentioned at the begining of section 3. For degrees 11 and 13 we have also succeeded to compute the necessary commuting extensions using the $S(Q, \Lambda_i)$ algorithm on which section 3 focused. To take into account the zero blocks of the commuting extensions an extra term

$$\sum_{i=1}^{d} \sum_{a=1}^{\dim \mathcal{P}_{q-1}} \sum_{b=n+1}^{N} \left((\tilde{Q} \Lambda_i \tilde{Q}^T)_{ab} \right)^2 \tag{90}$$

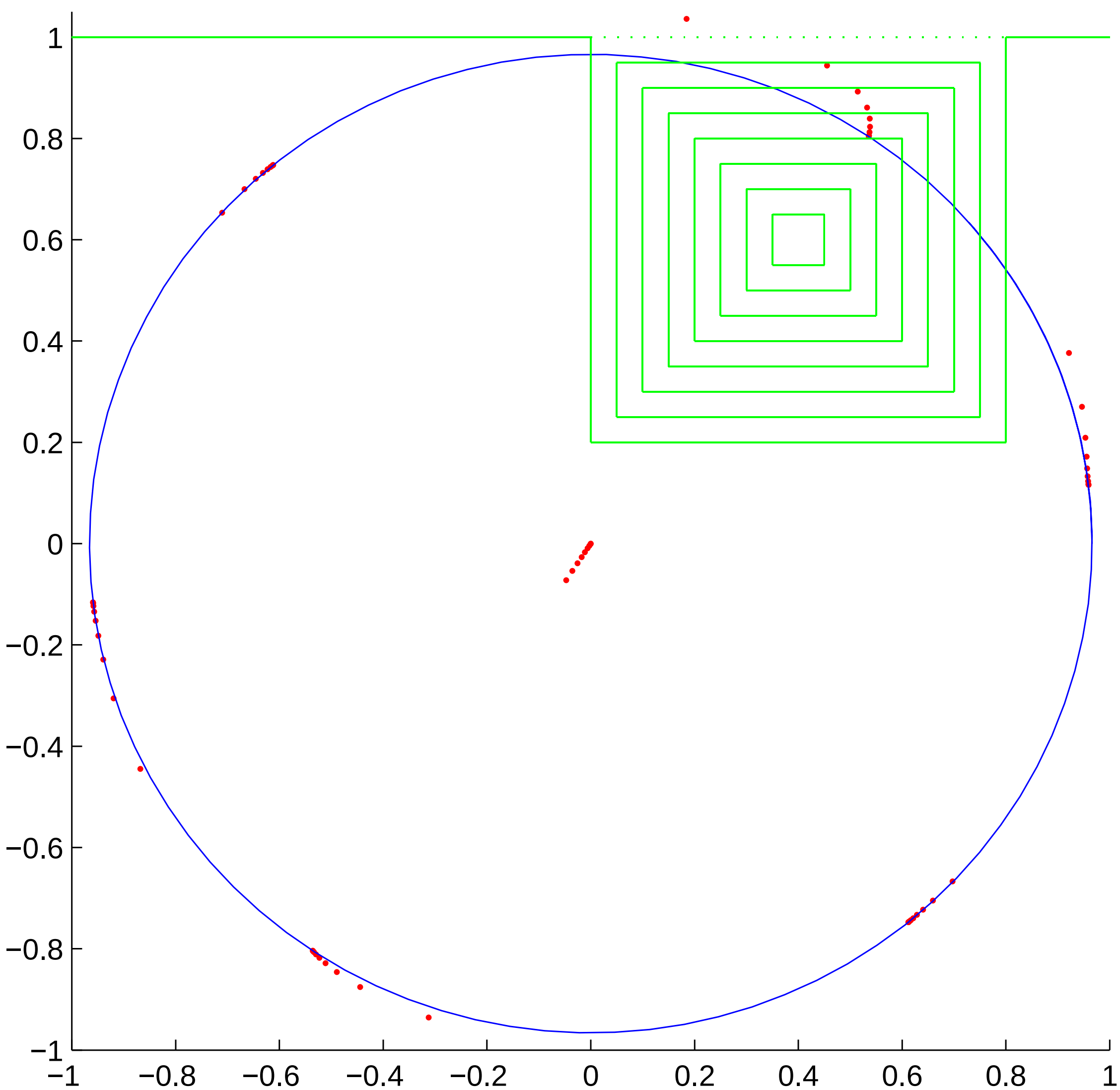

Figure 1: Radon formulae on a square with a smaller square removed, with uniform weight. As the size of the smaller square tends to zero, one node tends to the origin and 6 to the circle $x_1^2 + x_2^2 = \frac{14}{15}$. As the size of the smaller square is increased, one node moves outside the domain of integration, and eventually out of the larger square.

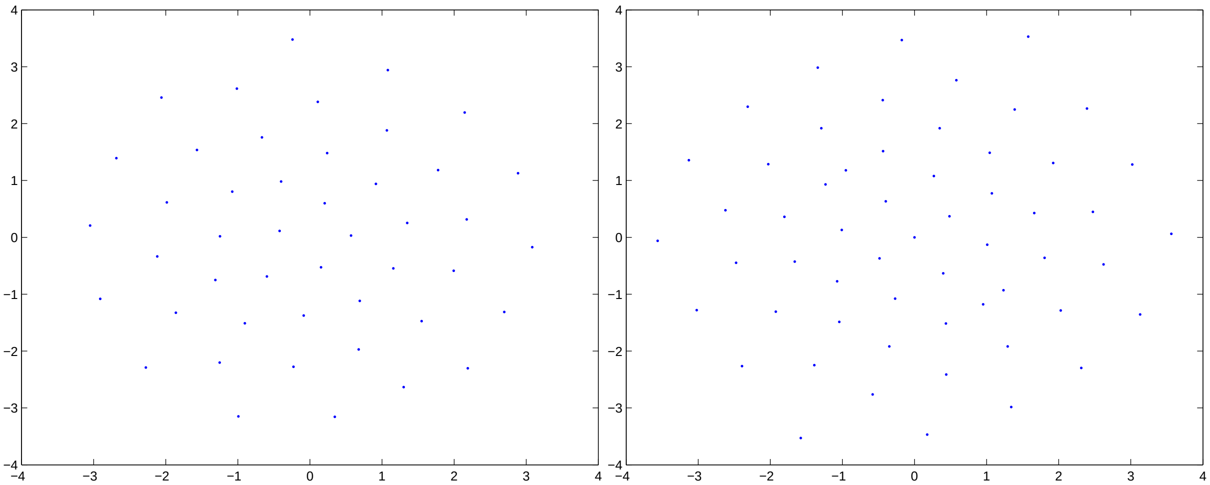

Figure 2: Nodes for a 46 point, degree 15 (left) and a 57 point, degree 17 (right) cubature formula for the plane $\Omega = \mathbf{R}^2$ with Gaussian weight $w(x_1, x_2) = e^{-x_1^2 - x_2^2}$

was added to S of (32), this gave a substantial improvement in performance (compared to the same algorithm with S not including the extra term). The nodes and weights are available at

`http://www.math.biu.ac.il/∼schiff/commext.html.`

Figure 2 displays the location of the nodes in the degree 15 and 17 formulae. The formulae found are not necessarilly minimal, and are certainly not unique, because of rotational symmetry. It seems our degree 13, 15 and 17 formulae are new, the one with order 17 having a smaller number of nodes compared to existing formulae of the same order, see [21].

7 Summary and open questions

The central results of this paper are theorems 9,10,11, which prove the equivalence of cubature formulae and commuting extensions satisfying the compatibility condition (equivalent in an appropriate basis to requiring certain zero blocks in the extension matrices). This raises the questions of existence and methods of computation for commuting extensions. Our knowledge of the theory of commuting extensions is summarized by theorems 1 to 6, and in section 3 we have described our initial attempts at their computation.

There is clearly enormous potential for further work here. In the context of our main topic, the connection between cubature formulae and commuting extensions, there is one nagging question that we have indicated several times in section 5: The vector space V on which the commuting extensions act does not yet have an interpretation as a space of functions (or maybe even polynomials). For numerical work in quantum mechanics it would

be a major advantage if we could contstruct finite dimensional function spaces containing the space of all polynomials of a given degree as a subspace, on which the natural projections of the operators x_i commute. The existence (or nonexistence) of such spaces is a topic we hope to investigate, see also [5].

Another question left open in our work is that we have not given an existence proof of cubature formulae from the commuting extension viewpoint. Although theorem 1 in section 2 guarantees the existence of commuting extensions of an arbitrary set of matrices, it does not guarantee extensions in the form we need to be able to apply theorem 11. An existence proof for commuting extensions of the required form would provide an alternative approach to *Tchakaloff's theorem* [22] that guarantees (for any suitable domain Ω and weight function $w(x)$) the existence of positive weight cubature rules that are exact for certain sets of functions.

The numerical question of computing cubature formulae is now subsumed under the more general question of computing commuting extensions, likewise the open sore that there is no good way to predict the minimal number of points needed for a cubature rule is subsumed under the question of finding the minimal dimension for commuting extensions. There are a number of points in the theory of commuting extensions which we feel may be improved, for example, existence of symmetric commuting extensions for symmetric matrices may well be provable, but we suspect the question of minimal dimension is extremely difficult. Fortunately, just because it is difficult theoretically does not mean answers cannot be found numerically, and we are hopeful that good algorithms can be devised that find commuting extensions of a given dimension, if they exist. The determining equations are linear and quadratic, and although there surely will be ill-conditioning in certain cases, it is hard to see why this should be so in general. We suspect that the poor performance of the $S(Q, \Lambda_i)$ algorithm in section 3 is to do with the fact we inforced equation (34), and only searched over Q. Likewise gradient and Newton algorithms possibly give exaggerated importance to linear components of the system. There is much more that can be tried here.

Another aspect to be considered in construction of cubature formulae/commuting extensions is symmetry in the domain Ω and weight $w(x)$. This will clearly influence the matrices A_i, which represent the natural projections of the operators x_i, and should be respected in construction of the extensions $\tilde{A}_i$, see also [5].

We hope very much that more applications will emerge for the notion of commuting extensions. The idea that noncommutativity can be resolved by introducing extra dimensions is a very natural one. In fact, we suspect that, more than the ranks or the norms of commutators, the size of minimal commuting extensions is probably the best measure of

how noncommuting a set of matrices is.

The minimal size issue appears in other settings too. For example, given a set of $m \times n$ matrices A_i we can ask what is the smallest N such that there exist an $m \times N$ matrix U and an $n \times N$ matrix V, both with orthonormal rows, and $N \times N$ diagonal matrices Λ_i, such that $A_i = U \Lambda_i V^T$. In our context, this provides a natural generalization of the singular value decomposition of a single matrix, in the same way that (26) provides the generalization of diagonalization of a single symmetric matrix.

Acknowledgements: We thank Jonathan Beck, Harry Dym, Yair Goldfarb, Gene Golub, Chen Greif, Ronny Hadani, David Kessler and Steve Shnider for suggestions and comments at various stages of this work.

References

[1] P.J.Davis and P.Rabinowitz, *Methods of Numerical Integration*, Academic Press (1984).

[2] R.Cools, *Constructing cubature formulae: the science behind the art*, Acta Numerica (1997) 1-54.

[3] G.H.Golub and J.H.Welsch, *Calculation of Gauss quadrature rules*, Math. Comp. **23** (1969) 221-230; W.Gautschi, *On the construction of Gaussian quadrature rules from modified moments*, Math.Comp. **24** (1970) 245-260; R.A.Sack and A.F.Donovan, *An algorithm for Gaussian quadrature given modified moments*, Numer.Math. **18** (1972) 465-478.

[4] J.Schiff and I.Degani, *Gaussian quadrature without orthogonal polynomials*, preprint.

[5] M.C.Cargo and R.G.Littlejohn, *Multidimensional DVRs from incomplete quadrature formulas*, *CIMMS-IPAM Workshop, Molecular Modeling and Computation: Perspectives and Challenges, November 2002*, `http://www.cimms.caltech.edu/workshops.html`.

[6] H.M.Möller, *Kubaturformeln mit minimaler Knotenzahl*, Numer.Math. **25** (1976) 185-200.

[7] A.S.Dickinson and P.R.Certain, *Calculation of matrix elements for one-dimensional quantum mechanical problems*, J.Chem.Phys. **49** (1968) 4209-4211; D.O.Harris, G.G. Engerholm and W.D. Gwinn, *Calculation of matrix elements for one-dimensional quantum mechanical problems and the application to anharmonic oscillators*, J.Chem.Phys. **43** (1965) 151.

[8] D.J.Tannor, *Introduction to Quantum Mechanics: A Time Dependent Perspective,* University Science Press, Sausalito (2004); J.C.Light, *Discrete-variable representations and their utilization*, in: I.Prigogine and S.A.Rice (Ed.), *Advances in chemical physics* **114** (2000).

[9] R.Dawes and T.Carrington Jr, *A multidimensional discrete variable representation basis obtained by simultaneous diagonalization*, *J.Chem.Phys* **121** (2004) 726-736.

[10] R.G.Littlejohn, M.Cargo, T.Carrington Jr, K.A.Mitchell and B.Poirier, *A general framework for discrete variable representation basis sets*, *J.Chem.Phys.* **116** (2002) 8691-8703.

[11] Eric W. Weisstein, *Sturmian Separation Theorem*, from *MathWorld - A Wolfram web resource*, `http://mathworld.wolfram.com/SturmianSeparationTheorem.html`.

[12] I.Degani, PhD Thesis, to be submitted to the Feinberg Graduate School, Weizmann Institute of Science in August 2004.

[13] P.Rosenthal, *Are almost commuting matrices near commuting matrices? Amer. Math. Soc. Monthly*, **76** (1969) 925-926.

[14] H.Lin, *Almost commuting selfadjoint matrices and applications*, *Fields Inst. Comm.* **13** (1997) 193-233.

[15] A. Bunse-Gerstner, R. Byers and V. Mehrmann, *Numerical methods for simultaneous diagonalization*, *SIAM J.Matrix Anal. and Appl.* **14** (1993) 927-949; J-F.Cardoso and A.Souloumiac, *Jacobi angles for simultaneous diagonalization*, *SIAM J. Matrix Anal. and Appl.* **17** (1996) 161-164.

[16] D.T.Pham, *Joint approximate diagonalisation of positive definite Hermitian matrices*, *SIAM J.Matrix Anal. and Appl.* **22** (2001) 1136-1152.

[17] R.Cools, I.P.Mysovskikh and H.J.Schmid, *Cubature formulae and orthogonal polynomials*, *J.Comp.Appl.Math.* **127** (2001) 121-152.

[18] H.M.Möller, *Lower Bounds for the number of nodes in cubature formulae*, in: G. Hämmerlin (Ed.) *Numerische Integration*, International Series of Numerical Mathematics, **45**, Birkhäuser, Basel (1979) 221-230.

[19] Y.Xu, *Block Jacobi matrices and zeros of multivariate orthogonal polynomials*, *Trans. Amer. Math. Soc.* **342** (1994) 855-866.

[20] J.Radon, *Zur mechanischen Kubatur*, Monat. fur Math. **52** (1948) 286-300.

[21] R.Cools, *An encylcopaedia of cubature formulas*, *J.Complexity*, **19** (2003) 445-453, `http://www.cs.kuleuven.ac.be/~nines/research/ecf/ecf.html`.

[22] V.Tchakaloff, *Formules de cubatures méchaniques à coefficients non negatifs*, *Bull. Sci. Math.* **81** (1957) 123-134; P.J.Davis, *A construction of nonnegative approximate quadratures*, *Math. Comp.* **21** (1967) 578-582.